PATENT
ATTY DOCKET NO. 5917/FET/FET

Control of Chemical Mechanical Polishing Pad Conditioner Directional Velocity to Improve Pad Life

Inventor: Young Joseph Paik

Related Applications

Ins. A1

This application is a continuation in part application of and claims priority from co-pending application serial number 60/298,878 filed June 19, 2001 and entitled "Advanced Process Control for Semiconductor Manufacturing Process."

This application is a continuation in part application of and claims priority from co-pending application serial number 60/305,798, filed July 16, 2001 and entitled "Feedforward and Feedback Control for Conditioning of Chemical Mechanical Polishing Pad."

This application is a continuation in part application of and claims priority from co-pending application serial number 60/318,743, filed September 12, 2001 and entitled "Feedforward and Feedback Control for Conditioning of Chemical Mechanical Polishing Pad."

Field of the Invention

The present invention is generally directed to the area of polishing and methods for improving the life and effectiveness of polishing pads in a chemical mechanical polishing process.

Background of the Invention

Chemical-mechanical polishing (CMP) is used in semiconductor fabrication processes for obtaining full planarization of a semiconductor wafer. The method involves removing material (e.g., a sacrificial layer of surface material) from the wafer, (typically silicon dioxide (SiO_2)) using mechanical contact and chemical erosion from, e.g., a moving polishing pad saturated with slurry. Polishing flattens out height differences, since areas of

high topography (hills) are removed faster than areas of low topography (valleys). FIG. 1A shows a top view of a CMP machine **100,** and FIG. 1B shows a side section view of the CMP machine **100** taken through line AA. The CMP machine **100** is fed wafers to be polished. Typically, the CMP machine **100** picks up a wafer **105** with an arm **101** and places it onto a rotating polishing pad **102**. The polishing pad **102** is made of a resilient material and is often textured, to aid the polishing process. The polishing pad **102** rotates on a platen **104** or turn table located beneath the polishing pad **102** at a predetermined speed. The wafer **105** is held in place on the polishing pad **102** by the arm **101**. The lower surface of the wafer **105** rests against the polishing pad **102**. The upper surface of the wafer **105** is against the lower surface of the wafer carrier **106** of arm **101**. As the polishing pad **102** rotates, the arm **101** rotates the wafer **105** at a predetermined rate. The arm **101** forces the wafer **105** against the polishing pad **102** with a predetermined amount of down force. The CMP machine **100** also includes a slurry dispense arm **107** extending across the radius of the polishing pad **102**. The slurry dispense arm **107** dispenses a flow of slurry onto the polishing pad **102**.

Over time the polishing pad loses its roughness and elasticity, and thus, its ability to maintain desired removal rates (polishing rates). It is known that the material removal rate provided by a given polishing pad decreases exponentially with time in the manner shown in FIG. 2. Further the decreased removal rate requires ever-increasing conditioning parameters, e.g., down force and/or conditioning angular velocity and/or conditioning time, in order to restore the desired removal rate of material from the wafer. As a consequence, the polishing pad must be conditioned (e.g., using a conditioning disk **108**), between polishing cycles. The conditioning disk is held in place on the polishing pad by arm **109**. As the polishing pad rotates, the conditioning disk **108** also rotates. Doing so roughens the surface of the pad and restores, at least temporarily, its original material removal rate. Furthermore, excessive pad conditioning shortens pad life.

A problem with conventional conditioning methods is that they may over-condition, e.g., wear out prematurely, the polishing pad. Each time a pad is replaced, one to several wafers must be polished thereon and the results measured, to ensure that the tool will yield the required polishing. This translates into processing delays and lost tool efficiency.

In an attempt to extend the life of the pad, one may selectively condition portions a polishing pad, or vary the down force of the conditioning element (e.g., conditioning disk **108**) along the surface of the CMP pad, based upon the distribution of waste matter across the planarizing surface. Other methods of extending pad life include varying the conditioning recipe across the surface of the polishing pad in response to polishing pad non-uniformities. However, these reported CMP processes are typically more concerned with improving the CMP process, e.g., improving within water non-uniformity, than in extending pad life.

Methods and devices that would extend pad life and therefore reduce the frequency of pad replacement offer significant cost savings to the wafer fabrication process.

Summary of the Invention

The present invention relates to a method, system and medium for conditioning a planarizing surface of a polishing pad in order to extend the working life of the pad. More specifically, at least some embodiments of the present invention use physical and/or chemical models (which can be implemented as a single model or multiple models) of the pad wear and wafer planarization processes to predict polishing pad performance and to extend pad life. This results in an increase in the number of semiconductor wafer or other substrates that can be polished with a single polishing pad, thereby providing significant cost savings in the CMP process, both in extending pad life and reducing the time devoted to pad replacement.

The model predicts polishing effectiveness (wafer material removal rate) based on the "conditioning" operating parameters of the conditioning process. In at least some embodiments of the present invention, operating parameters of conditioning include angular direction and angular velocity of a conditioning disk (where a disk is used to condition) during conditioning, and may also include other factors, such as the frequency of conditioning and time of conditioning. The model selects, and then maintains, polishing pad conditioning parameters within a range that does not overcondition the pad while providing acceptable wafer material removal rates. Thus the present invention provides a process for the feed forward and feed backward control of the CMP polishing process.

In one aspect of the invention, a method of conditioning a planarizing surface in a CMP apparatus having a polishing pad and a conditioning disk includes polishing a wafer in the CMP apparatus under a first set of pad conditioning parameters selected to maintain wafer material removal rates within preselected minimum and maximum removal rates; measuring a wafer material removal rate occurring during said polishing step; calculating, based upon said wafer material removal rate, updated pad conditioning parameters to maintain wafer material removal rates within the maximum and minimum removal rates; and conditioning the polishing pad using the updated pad conditioning parameters. The updated pad conditioning parameters are calculated using a pad wear and pad recovery model by calculating wafer material removal rate as a function of pad conditioning parameters including conditioning disk rotational speed and direction; and determining the difference between the calculated and measured wafer material removal rates and calculating updated pad conditioning parameters to reduce said difference, wherein the updated conditioning parameters are updated according to the equation $k = (k_1) + g * (k - (k_1))$, where k is a measured parameter, k_1 is calculated parameter estimate, g is the estimate gain and $(k-(k_1))$ is the prediction error.

In at least some embodiments of the invention, the first set of pad conditioning parameters are determined empirically, or using historical data, or using the results of the design of experiment (DOE).

In at least some embodiments of the invention, the pad conditioning parameters of the pad wear and pad recovery model further includes frequency of conditioning, or time of conditioning, or translational speed of conditioning disk during conditioning.

In at least some embodiments of the invention, the step of determining the wafer material removal rate includes measuring the wafer thickness before and after polishing.

In at least some embodiments of the invention, the step of calculating updated pad conditioning parameters in step (c) includes executing a recursive optimization process, or in at least some embodiments, includes calculating conditioning parameters such that the parameter is within determined maximum and minimum values.

In at least some embodiments of the present invention, the gain is an indication of variability or reliability in the measured parameter, and the gain is in the range of about 0.5 to 1.0, or gain is in the range of about 0.7 to 0.9.

In at least some embodiments, updated pad conditioning parameters are calculated by determining a difference between an output of the pad wear and pad conditioning model and the wafer material removal step (c). In at least some embodiments, this difference is minimized.

In at least some embodiments of the invention, the steps (b) through (e) are repeated.

In at least some embodiments of the invention, the maximum value for wafer material removal rate is the saturation point of the wafer material removal rate vs. conditioning down force curve, or in at least some embodiments, the maximum value for wafer material removal rate is the initial rate, or in at least some embodiments, the minimum value for wafer material removal rate is defined by the maximum acceptable wafer polishing time.

In at least some embodiments of the invention, the wafer material removal rate is defined by the equation

$$\mathrm{Re}\,movalRate\Big]_{\min}^{\max} = f\left(\omega_{disk}\Big]_{\min}^{\max}, f\Big]_{\min}^{\max}, t_{conditioning}\Big]_{\min}^{\max}, direction, T_2\Big]_{\min}^{\max}\right),$$

where ω_{disk} is the angular velocity of the conditioning disk during conditioning of the polishing pad, t is the time of conditioning, f is the frequency of condition, direction is the spinning direction of the conditioning disk, and T_2 is the sweeping speed of the conditioning disk during conditioning.

In another aspect of the invention, an apparatus for conditioning polishing pads used to planarize substrates includes a carrier assembly having an arm positionable over a planarizing surface of a polishing pad; a conditioning disk attached to the carrier assembly; and an actuator capable of controlling an operating parameter of the conditioning disk; and a controller operatively coupled to the actuator, the controller operating the actuator to adjust the operating parameter of the conditioning disk as a function of a pad wear and pad recovery model that predicts the wafer material removal rate based upon conditioning pad rotational speed and direction.

In at least some embodiments of the invention, the updated pad conditioning parameters are calculated using a pad wear and pad recovery model by calculating wafer material removal rate as a function of pad conditioning parameters including conditioning disk rotational speed and direction; and determining the difference between the calculated and measured wafer material removal rates and calculating updated pad conditioning parameters to reduce said difference, wherein the updated conditioning parameters are updated according to the equation $k = (k_1) + g * (k - (k_1))$, where k is a measured parameter, k_1 is calculated parameter estimate, g is the estimate gain and $(k-(k_1))$ is the prediction error.

In at least some embodiments, the pad conditioning parameters of the pad wear and pad recovery model further includes frequency of conditioning, time of conditioning, or speed of conditioning disk during conditioning.

In at least some other embodiments of the present invention, the gain is an indication of variability or reliability in the measured parameter.

In another aspect of the invention, a method of developing a pad wear and pad conditioning model for optimization of the pad conditioning for polishing pads used to remove material from a wafer, is provided. The method includes:

a) determining the relationship between at least one pad conditioning parameter and wafer material removal rate;

b) determining maximum and minimum values for each of the at least one pad conditioning parameters and the wafer material removal rate; and

c) recording the relationships and minimum and maximum values of the at least one pad conditioning parameter and the wafer removal rate.

In at least some embodiments of the invention, the at least one pad conditioning parameter includes a plurality of parameters and the wafer removal rate is defined as a weighted function of the plurality of pad conditioning parameters, or in at least some embodiments, the at least one pad conditioning parameters includes conditioning disk rotational speed, or in at least one embodiment, the at least one pad conditioning parameter further includes conditioning disk rotational direction.

In at least some embodiments of the invention, the at least one pad conditioning parameter includes one or more parameters selected from the group consisting of conditioning disk down force, conditioning disk rotational rate and direction, frequency and duration of conditioning, and conditioning disk translational speed.

In at least some embodiments of the invention, the relationship between the at least one conditioning parameter and wafer removal rate is determined by incrementally varying the conditioning parameter and measuring the resultant wafer removal rate.

In at least some embodiments of the invention, the maximum value for the conditioning parameter is the value above which no incremental increase of the wafer removal rate is observed, or in at least some embodiments, the minimum value for the conditioning parameter is the value which provides the minimum wafer removal rate.

In at least some embodiments of the invention, the invention further includes polishing a wafer in the CMP apparatus under a first set of pad conditioning parameters selected to maintain wafer material removal rates within preselected minimum and maximum removal rates including conditioning disk rotational speed and direction, determining a wafer material removal rate occurring during said polishing step, calculating updated pad conditioning parameters based upon said determined wafer material removal rate and the pad wear and conditioning model to maintain wafer material removal rates within the maximum and minimum removal rates, and conditioning the polishing pad using the updated pad conditioning parameters.

In at least some embodiments of the invention, the updated pad conditioning parameters are calculated by determining the difference between an output of the pad wear and conditioning model and said determined wafer material removal, or in at least some embodiments, the updated pad conditioning parameters are updated according to the equation $k = (k-1) + g*(k-(k-1))$, where k is a measured wafer material removal rate, k_1 is a calculated wafer material removal rate, g is the estimate gain, and $(k-(k_1))$ is the prediction error.

In another aspect of the invention, a computer readable medium is provided having instructions being executed by a computer, the instructions including a computer-implemented software application for a chemical mechanical polishing process. The instructions for implementing the process include:

a) receiving data from a chemical mechanical polishing tool relating to the wafer removal rate of at least one wafer processed in the chemical mechanical polishing process; and

b) calculating, from the data of step (a), updated pad conditioning parameters within defined maximum and minimum values, wherein the updated pad conditioning parameters are calculated by determining the difference between an output of a pad wear and conditioning model and the data of step (a).

In at least some embodiments of the invention, calculating updated conditioning parameters includes calculating parameters such that the parameter is within the determined minimum and maximum values, or in at least some embodiments, calculating updated pad conditioning parameters in step (b) comprises executing a recursive optimization process.

In at least some embodiments of the invention, the maximum value for wafer material removal rate is the saturation point of the wafer material removal rate vs. conditioning down force curve, or in at least some embodiments, the maximum value for wafer material removal rate is the initial rate, or in at least some embodiments, the minimum value for wafer material removal rate is defined by the minimum acceptable wafer polishing time.

In at least some embodiments of the invention, the difference is adjusted using an estimate gain prior to calculating updated pad conditioning parameters.

In another aspect of the invention, a method of conditioning a planarizing surface in a chemical mechanical polishing (CMP) apparatus having a polishing pad against which a wafer is positioned for removal of material therefrom and a conditioning disk is positioned for conditioning of the polishing pad is provided. The method includes:

(a) developing a pad wear and pad conditioning model that defines wafer material removal rate as a function of pad conditioning parameters by:

(i) determining the relationship between at least one pad conditioning parameter and wafer material removal rate;

(ii) determining maximum and minimum values for each of the at least one pad conditioning parameters and the wafer material removal rate;

(iii) recording the relationships and minimum and maximum values of the at least one pad conditioning parameter and the wafer removal rate;

(b) polishing a wafer in the CMP apparatus under a first set of pad conditioning parameters including conditioning disk rotational speed and direction, selected to maintain wafer material removal rates within preselected minimum and maximum removal rates;

(c) determining a wafer material removal rate occurring during said polishing step;

(d) calculating updated pad conditioning parameters based upon said determined wafer material removal rate of said step (b) and the pad wear and conditioning model to maintain wafer material removal rates within the maximum and minimum removal rates, and

(f) conditioning the polishing pad using the updated conditioning parameters.

In another aspect of the invention, a system for conditioning a planarizing surface in a chemical mechanical polishing (CMP) apparatus having a polishing pad against which a wafer is positioned for removal of material therefrom and a conditioning disk is positioned for conditioning of the polishing pad includes:

a) a pad wear and conditioning model that defines wafer material removal rate as a function of at least one pad conditioning parameters including rotation and direction of the conditioning disk;

b) polishing means for polishing a wafer in the CMP apparatus

c) measuring means for determining a wafer material removal rate; and

d) calculating means for updating pad conditioning parameters based upon a wafer material removal rate measured using means of step (c) and the pad wear and conditioning model to maintain wafer material removal rates within the maximum and minimum removal rates.

Thus, polishing pad life is extended by using an appropriate conditioning angular velocity to keep within the acceptable range of wafer material removal rate and reversing direction of conditioning and/or adjusting angular velocity or other conditioning parameters whenever the removal rate drops below the acceptable removal rate. By applying a "one size fits all" approach to pad conditioning parameters, e.g., by determining conditioning parameters without accounting for actual change in wafer material removal rates, conventional processes overcompensate, thereby removing more pad material than is necessary and accelerating pad wear. In contrast, the present invention thus provides improved conditioning parameters by determining only those forces that are minimally necessary to recondition the damaged pad.

Brief Description of the Figures

Various objects, features, and advantages of the present invention can be more fully appreciated with reference to the following detailed description of the invention when considered in connection with the following drawings.

Figures 1A-B show a CMP machine. Figure 1A shows a top plan view of a conventional CMP machine. Figure 1B shows a side sectional view of the conventional CMP machine from Figure 1A taken through line A--A.

Figure 2 shows an example exponential decay of wafer material removal rate and the equilibrium state of the removal rate that occurs between Figures 3B and 3C.

Figures 3A-C are cross-sectional views of polishing pads. Figure 3A is a view of a new polishing pad. Figure 3B is a view of a polishing pad after a single (or few) conditioning event. Figure 3C shows an old polishing pad whose surface asperities have been smoothed out by repeated conditioning.

Figures 4A-C are cross-sectional views of polishing pads. Figure 4A is a view of a new polishing pad. Figure 4B is a view of a polishing pad after conditioning in a first angular direction. Figure 4C shows the same polishing pad after conditioning in the opposite angular direction.

Figure 5 is a flow diagram of the feedback loop used in at least some embodiments of a CMP process optimization.

Figure 6 is a flow diagram illustrating an example of data collection and generation of a pad wear and conditioning model.

Figure 7 is a model of polishing pad wear based on Figures 3 and 4 used in predicting and optimizing the water removal rate in a CMP process.

Figure 8 is a model of polishing pad recovery based on Figures 3 and 4 used in predicting and optimizing the water removal rate in a CMP process.

Figure 9 is a model based on Figures 5 and 6 for predicting and modifying CMP operating parameters to optimize the wafer process.

Figure 10 is a side sectional view of a CMP machine for use in at least some embodiments of the present invention.

Figure 11 is a block diagram of a computer system that includes tool representation and access control for use in at least some embodiments of the invention.

Figure 12 is an illustration of a floppy disk that may store various portions of the software according to at least some embodiments of the invention.

Detailed Description of the Invention

Novel methods for feed forward and feed back controls of the CMP process for maximizing the life of the polishing pad are described herein. Extended pad life results in reduced down time for the CMP process because the polishing pad can polish more wafers over a longer period of time without requiring replacement or adjustment (e.g., removal of the damaged portion of the pad). The term wafer is used in a general sense to include any substantially planar object that is subject to polishing. Wafers include, in additional to monolith structures, substrates having one or more layers or thin films or other architecture deposited thereon.

The polishing pad surface needs to maintain a certain level of roughness and elasticity in order to provide the required wafer material removal rates in a CMP process. The roughness and elasticity of the pad decreases with successive wafer polishes, thereby reducing the wafer material removal rate. Initial polishing pad surface conditions (asperities **320**) are shown in FIG. 3A, at which time surface roughness is at a maximum. After the pad has been used for polishing, these asperities are pushed down, often in varying directions. To compensate for this, and restore at least some of the roughness of the pad, the pad is conditioned using, for example, a conditioning disk that rotates, for example, in the direction indicated by arrow **340** shown in FIG. 3B. Although the invention is described herein with disk style conditioners, the use of other conditioning mechanisms is specifically contemplated. This, however, introduces a directional bias into the pad surface features **320**. Subsequent conditioning operations using the same direction of conditioning may lead to increased directionality in pad surface apserities, thereby blocking the flow of the slurry in the pad and also reducing the contact surface between the pad asperities and the polishing wafer. This is shown by the even greater directional bias of the asperities **320** of FIG. 3C. As a result, wafer material removal rates worsen as directional bias of the pad surface features increases. FIG. 2 shows the decrease in removal rate over time as a result of the conditioning process shown in FIGs 3A-C.

FIGs. 4A, 4B and 4C represent a simplified model used for overcoming the aforementioned bias issue, wherein the angular velocity of the conditioning disk is alternated. Referring first to FIG. 4A, this figure shows initial polishing pad surface conditions. The polishing pad **400** is conditioned by contacting the pad with a conditioning disk at a first angular velocity (e.g., clockwise, indicated by arrow **420** in FIG. 4B), which introduces some directionality to the polishing pad surface features **440**. In a subsequent conditioning event, the angular velocity of the conditioning disk is reversed (e.g., counterclockwise, as shown by arrow **460** in FIG. 4C) to "undue" the effect of the previous conditioning events. Alternating the speed and direction of conditioning extends the surface roughness and elasticity. The process shown in FIGs. 4A, 4B and 4C may be repeated for the entire life cycle of the pad until no more active sites are available.

Thus, the polishing pad may be conditioned in a first direction for a predetermined number of times after which the direction of conditioning is reversed. The optimal number of conditioning events in a particular direction is expected to change (decrease) as the pad ages. The model for pad conditioning and recovery adjusts the process accordingly.

The mechanical processes described above during wafer planarization and conditioning of the polishing pad provide a model for optimization of the planarization process. By adjusting pad conditioning parameters according to this model, the pad life can be extended without compromise to the wafer material removal rate. In particular, speed and direction of the conditioning disk, an optionally other operating variables such as conditioning frequency, conditioning duration, and transitional speed of conditioning disk across the pad surface, are adjusted in a feed forward and feed back loop that predicts and then optimizes pad conditioning operating parameters.

According to at least one embodiment of the present invention, an initial model is developed based upon knowledge of the wafer polishing process, and is used in at least some embodiments of the present invention as is shown in FIG. 5. Based on that initial model, e.g., the wafer and polishing pad parameters remain constant, initial processing conditions are identified that will provide a wafer material removal rate between a preselected minimum and maximum value for a given set of conditioning parameters, hereinafter, the "acceptable" range for wafer material removal rates. The conditions are selected to prevent overconditioning of the pad.

Referring now to FIG 5, wafers are polished according to the initial conditions in step **500**. The thicknesses of the polished wafers are measured and a wafer material removal rate is calculated in step **510**, which information is then used in a feedback loop to maintain the wafer material removal rate within the accepted range. The actual removal rate is compared with the predicted removal rate (derived from the pad wear model). Deviations, i.e., prediction errors, are used to adjust pad conditioning parameters in step **520** according to the model of the invention to compensate for the reduced polishing capability of the polishing pad as identified in the model and/or to correct for any un-modeled effects. The polishing pad is conditioned according to the updated conditioning parameters in step **530**. Polishing is

repeated in step **540** and the polishing results are used to further update the polishing conditions by repeating steps **510-530**.

By maintaining the wafer material removal rate and conditioning parameters within the preselected minimum and maximum range, overconditioning of the pad is prevented; that is, conditioning parameters may be used that are just sufficient to restore polishing pad effectiveness, but which do not unduly damage the pad. In operation, it may be desirable to select pad conditioning parameters that result in wafer material removal rates that are close to the minimum acceptable rates, as these conditioning forces are less aggressive and therefore are more likely to avoid overconditioning of the polishing pad. However, one should be cautious (or at least cognizant) about operating too closely to the minimum removal rate since a sudden degradation in the pad condition may cause the wafer material removal rate to drop below the minimum acceptable rate.

Pad conditioning optimization is carried out with reference to a specific polishing system. That is, the conditions that improve pad lifetime are specific to the type of wafer being polished, the slurry used in polishing and the composition of the polishing pad. Once a wafer/slurry/polishing pad system is identified, the system is characterized using the models developed and discussed herein. Exemplary polishing pad and wafer parameters include polishing pad size, polishing pad composition, slurry composition, wafer composition, rotational velocity of the polishing pad, polishing pad pressure, and translational velocity of the wafer.

In at least some embodiments of the present invention, it is envisioned that a separate model (or at least a supplement to a composite model) is created for each slurry/polishing pad wafer combination (i.e., for each different type/brand of slurry and each type/brand of pad) that may be used in production with a given type of wafer.

FIG. 6 shows a flow diagram of the steps used in the development of the pad wear and conditioning model in at least some embodiments of the invention. In the design of experiment (DOE) in step **600**, that is, a set of experiments used to define the model, the relationship between wafer material removal rate and a first conditioning parameter x_1, e.g., conditioning disk angular velocity (rpm), is determined using the selected polishing system.

The relationship is determined by measuring wafer material removal rates at different conditioning disk angular velocities with wafer parameters such as polishing force, polishing duration, etc., held constant. Thus, a wafer is polished under specified conditions, e.g., for a specified time and at specified polishing pad and wafer speeds, and the rate of material removal is determined. Pad conditioning and wafer polishing (the "polishing event") may be carried out simultaneously, i.e., using an apparatus such as shown in FIG. 10, or pad conditioning may be followed by wafer polishing. The conditioning disk velocity is increased incrementally from wafer to wafer (or thickness measurement to thickness measurement) with all other parameters held constant, and the wafer removal rate is again determined. A curve as shown in FIG. 7 may be generated, which illustrates the effect of the conditioning disk velocity on the wafer's material removal rate for a given polishing system (all other parameters being held constant). The curve is represented as a linear curve over the removal rate of interest, but may, in at least some embodiment of the invention, be a non-linear, e.g. quadratic or exponential curve.

In step **610** of FIG. 6, as contemplated by at least some of the embodiments of the invention, minimum and maximum values for the conditioning parameter are determined. With reference to FIG. 7, a curve **700** represents the change in wafer material removal rate with time (as determined by number of wafers polished) for a given set of operating parameters. The removal rate is represented as decreasing linearly with time until an equilibrium rate is achieved, which may be, but is not required to be, less than the minimum removal rate **770**, which is determined by the operator, for example, based upon production needs. The slope of the curve is characterized by the angle θ_1. The curve can be, in at least some of embodiments, linear or non-linear, e.g. exponential or quadratic, or the like. The minimum wafer material removal rate is dictated by production goals, since a minimal wafer throughput rate is needed. The maximum wafer material removal rate **795** is defined as the initial removal rate.

Successive curves **720, 740, 760** may also be generated for different conditioning disk velocities (here increasing velocities are shown). The removal rate range **780** defines the removal rate maximum and minimum for the model, where the maximum removal rate is the

initial removal rate and the minimum removal rate is production determined. Intersection of curves **700, 720, 740, 760** with the minimum removal rate defines the upper limit of wafers that can be polished under the defined pad conditioning parameters. The angles θ_1, θ_2, θ_3, and θ_4 are defined by the intersection of the equilibrium curve **790** with curves **700, 720, 740, 760**, respectively. The values for θ are descriptive of the response of the polishing process to conditioning parameters. The larger the value for θ, the steeper the slope of the curve and the more sensitive the planarization process is to conditioning parameters.

In step **620**, as contemplated by at least some embodiments of the present invention, the relationship between wafer material removal rate and a second conditioning parameter, e.g., direction of pad conditioning, is determined in the same polishing system. In step **630**, x_2, maximum and minimum values for the second conditioning parameter and wafer material removal rates is determined.

As is described above with reference to FIGs. 3 and 4, once the equilibrium wafer material removal rate or the minimum acceptable material removal rate is reached, recovery is possible by reversing the direction of pad conditioning (see, FIG. 4C). With reference to FIG. 8, a curve is generated to illustrate the effect of direction of conditioning pad rotation on wafer removal rate (all other variables held constant). Curve **800** represents the increase in wafer material removal rate with time (as determined by number of wafers polished) for a given angular velocity as the flattening of the polishing pad surface is alleviated upon conditioning in the reverse direction. The removal rate is shown as increasing linearly with time until a saturation point **810** is achieved, which is typically less than the initial maximum removal rate of the pad. In at least some embodiments of the invention, the curve may be linear or non-linear, e.g. expotential or quadratic, or the like. Additional polishing results in loss of surface roughness due to orientation in the opposite direction, and so wafer material removal rates again are expected to decline. Thus, the maximum wafer material removal rate **815** is defined at the curve maximum. As above, an operating minimum wafer material removal rate **825** can be suggested by production goals, since a minimal wafer throughput rate is needed. The removal rate range **880** defines the removal rate maximum and minimum for the pad recovery model.

In at least some embodiments of the invention, successive curves **820, 840, 860** are also generated for different velocities of the conditioning disk. Each curve reaches a maximum, which represents the optimal recovery of the polishing pad due to reversal of the conditioning direction and then declines. The angles φ_1, φ_2, φ_3, and φ_4 are defined for each curve **800, 820, 840, 860,** respectively. The value for φ is descriptive of the recovery of the polishing pad. The larger the value for φ, the steeper the slope of the curve and the more sensitive the recovery process is to conditioning parameters. Since it is not possible to fully compensate for pad wear by reversing direction of conditioning, for a given sample curve conditioned at a given angular velocity, $\theta > \varphi$.

According to the above model, once the maximal recovery in wafer material removal rates is achieved, the wafer material removal rate will again decline and approach the minimum acceptable removal rate. The direction of the conditioning disk is again reversed and wafer material removal rates are expected to increase once again. The values for each successive maximum in the recovery curves of FIG. 8 are expected to decrease until performance above the minimum removal rate is not possible. At this point, the conditioning velocity is increased in order to bring the removal rate into the acceptable range. The model at the higher velocity is now used to predict future performance.

The results of these studies provide maximum and minimum wafer material removal rates, and performance at different conditioning velocities. In addition, values for constants $\theta_1 - \theta_4$ and $\varphi_1 - \varphi_4$ relating to curve slopes may be determined. Although the above designs of experiment show a conditioning parameter that demonstrates an increase in wafer removal rate with increase in magnitude of the parameter, it is understood that the opposite relationship can exist, so that the minimal parameter value produces the maximum wafer removal rate. The models can be adjusted accordingly. Maximum and minimum conditions may be determined for any combination of polishing pad, wafer and polishing slurry known in the art. Additional parameters, up to x_n, may be included in the model and their minimum and maximum values determined as indicated by steps **640** and **650**.

The model can be represented as raw data that reflects the system, or it can be represented by equations, for example multiple input-multiple output linear, quadratic and

non-linear equations, which describe the relationship among the variables of the system. Feedback and feed forward control algorithms are constructed in step **660** based on the above model using various methods. For example, the wafer removal rate may be defined as the weighted contribution of all the pad conditioning parameters, x_1 through x_n. The algorithms may be used to optimize conditioning parameters using various methods, such as recursive parameter estimation. Recursive parameter estimation is used in situations such as these, where it is desirable to model on line at the same time as the input-output data is received. Recursive parameter estimation is well-suited for making decisions on-line, such as adaptive control or adaptive predictions. For more details about the algorithms and theories of identification, see Ljung L., *System Identification - Theory for the User, Prentice Hall*, Upper Saddle River, N.J. 2nd edition, 1999.

In at least some embodiments of the present invention, the CMP pad life is a function of surface roughness and pad elasticity as shown in eq. 1:

$$PadLife = f(surface\ roughness,\ elasticity). \qquad (1)$$

In at least some embodiments of the present invention, the wafer material removal rate is described according to eq. 2:

$$\mathrm{Re}\,movalRate\Big]_{\min}^{\max} = f(\omega_{disk}\Big]_{\min}^{\max}, f\Big]_{\min}^{\max}, t_{conditioing}\Big]_{\min}^{\max}, direction, T_2\Big]_{\min}^{\max}), \qquad (2)$$

where ω_{disk} is the angular velocity (rotational speed, e.g., rpm) of the conditioning disk during conditioning of the polishing pad, direction is the direction of spin, i.e., clockwise or counterclockwise, of the conditioning disk, T_2 is the translational speed of the conditioning disk across the pad surface, as shown in the exemplary CMP device in FIG. 10 (which will be described in greater detail below), $t_{conditioning}$ is the duration of conditioning, and f is frequency of conditioning. Frequency is measured as the interval, e.g., number of wafers polished, between conditioning events and direction is defined above. For example, a frequency of 1 means that the pad is conditioned after every wafer, while a frequency of 3 means that the pad is conditioned after every third wafer. The sweeping speed is the speed at which the conditioning disk moves across the surface of the polishing pad. The motion is

indicated by arrow T_2 in FIG. 10. For the purposes of initial investigation, it is assumed in at least some embodiments of the present invention that t (time), T_2 (sweep speed), and f (frequency) are held constant.

The objective function is to maintain removal rates within the minimum and maximum allowable rates (the "acceptable rates") by controlling the conditioning disk speed and direction, and, optionally, by controlling other factors such as frequency and duration of conditioning, conditioning disk down force, speed of translation of the conditioning disk across the pad surface. Each of the conditioning parameters is maintained within their determined boundary conditions, i.e., minimum and maximum values, as well.

The CMP parameters (variable) and constants from the model may then be programmed into a computer, which may then constantly monitor and appropriately vary the parameters during the process to improve the wafer material removal rate and the pad life, as shown in FIG. 9. Parameters from the base study **901** are input into the computer or other controller **902,** which runs the wafer polishing process, and the estimator **903,** which monitors and modifies the process parameters. The actual output (i.e., measured removal rate) **904** is monitored and compared to the predicted output (i.e., target removal rate) **905** calculated by estimator **903**. The difference **906** between the actual output **904** and the predicted output **905** is determined and reported **907** to the estimator **903**, which then appropriately generates updated parameters **908** for the process **902**.

Updating model parameters for feedback control is based on eq. 3.

$$k = (k_1) + g * (k - (k_1)), \tag{3}$$

where k is a current parameter, k_1 is a previous parameter estimate, g is the estimate gain and $(k-(k_1))$ is the prediction error. Estimate gain is a constant selected by the user, which is used as a measure of machine error or variability. Gain factor may be determined empirically or by using statistical methods. In at least some embodiments, the gain is in the range of about 0.5 to 1.0, or in at least some embodiments, in the range of about 0.7 to 0.9.

By way of example, a series of curves may be generated for a polishing system of interest as described above for determining the relationship between wafer material removal

rate and conditioning disk rotational velocity and direction. Curves are generated using a standard polishing procedure, with all operating conditions held constant with the exception of the parameter(s) under investigation. Exemplary polishing pad and wafer parameters that are held constant include polishing pad size, polishing pad composition, wafer composition, polishing time, polishing force, rotational velocity of the polishing pad, and rotational velocity of the wafer. The variable parameters include at least the angular speed and direction of the conditioning disk; however, additional parameters may be included in the model. Using the model such as shown in FIG. 6 for at least some of the embodiments of the invention, and the curves generated as in FIGs. 7 and 8, values for θ_1- θ_4, φ_1- φ_4, minimum and maximum values for wafer material removal rate, conditioning down force and conditioning disk rotational velocity are determined. An algorithm that models the wafer planarization is defined, and a first set of pad conditioning parameters may be determined for a polishing system of interest, either empirically or using historical data or data from the DOE.

An algorithm which models the pad wear and pad recovery process is input into the estimator and a predicted wafer material removal rate is calculated based upon the model. The actual results are compared against the predicted results and the error of prediction is fed back into the estimator to refine the model. New conditioning parameters are then determined. Based upon the models described herein, these parameters are just sufficient to revitalize the pad surface without overconditioning. Thus, the smallest increment in conditioning parameters that meet the model criteria is selected for the updated conditioning parameters. Subsequent evaluation of the updated model will determine how good is the fit, and further modifications can be made, if necessary, until the process is optimized.

In at least some embodiments of the present invention, the conditioning parameters are updated in discrete increments or steps, defined by way of example, by the incremental curves shown in FIGs. 7 and 8. A suitable number of curves are generated so that steps are small enough to permit minor adjustments to the conditioning parameters.

Also, in at least some embodiments of the present invention, the updated conditioning parameters may be determined by interpolation to the appropriate parameters, which may lie

between curves. Interpolation may be appropriate in those instances where a fewer number of curves are initially generated and the experimental results do not provide a fine resolution of the parameters.

While deviations from the predicted rate reflects, in part, the inability of the model to account for all factors contributing to the process (this may be improved with subsequent iterations of the feedback process), deviations from the predicted wafer material removal rate over time represent a degradation in CMP pad polishing. By identifying and modifying the pad conditioning process to account for these changes in polishing capabilities, optimal wafer material removal rates are maintained without overconditioning of the condition pads, e.g., operating above the saturation point of the system.

An additional feature of the method is the use of gain factor to qualify the prediction error, as shown in eq. 3. Thus, the method suggests that the model need not correct for 100% of the deviation from predicted value. A gain factor may be used to reflect uncertainty in the measured or calculated parameters, or to "damp" the effect of changing parameters too quickly or to a too great an extent. It is possible, for example, for the model to overcompensate for the prediction error, thereby necessitating another adjustment to react to the overcompensation. This leads to an optimization process that is jumpy and takes several iterations before the optimized conditions are realized. Use of a gain factor in updating the parameters for feedback control qualifies the extent to which the model will react to the prediction error.

Once the basic system is understood and optimized, it is possible to empirically vary other conditioning operating parameters and access their impact on pad conditioning and wafer polishing. For example, conditioning down force, which may be set to a constant value in the initial study, may be increased (or decreased). The system is monitored to determine the effect this change had on the system. It should be readily apparent that other factors relevant to pad wear and conditioning may be evaluated in this manner. By way of example, conditioning time (residence time of the disk on the pad), conditioning disk translational speed, conditioning down force, and the like may be investigated in this manner. In addition, the model may be modified to include additional parameters in the model.

It is envisioned that at least some embodiments of the present invention may be practiced using a device **1000** such as the one shown in FIG. 10. The apparatus has a conditioning system **1010** including a carrier assembly **1020**, a conditioning disk **1030** attached to the carrier assembly, and a controller **1040** operatively coupled to the carrier assembly to control the down force (F) and rotation rate (ω) of the conditioning disk. The carrier assembly may have an arm **1050** to which the conditioning disk **1030** is attached and means **1060a-d** to move the conditioning disk in and out of contact with the planarizing surface. For example, the controller **1040** may be operatively coupled to the moving means to adjust the height and position of the arm carrying the conditioning disk (**1060a, 1060b, 1060c, 1060d**). Similar controls for control of the position and movement of the wafer may also be present. In operation, the controller adjusts the operating parameters of the conditioning disk, e.g., down force and rotation rate, in response to changes in wafer material removal rate. The controller may be computer controlled to automatically provide conditioning according to the calculated conditioning recipe. Thus, the apparatus provides a means for selectively varying the pad conditioning parameters over the operating life of the pad **1080** in order to extend pad life without compromise to the planarization process of the wafer **1090**. Other types of devices where, e.g., other components have their height, positions, and/or rotations adjusted are also contemplated by at least some embodiments of the present invention.

Additional apparatus utilized to implement the feedforward and feedback loop include a film thickness measurement tool to provide thickness data needed to calculate wafer material removal rate. The tool may be positioned on the polishing apparatus so as to provide in-line, *in situ* measurements, or it may be located remote from the polishing apparatus. The tool may use optical, electrical, acoustic or mechanical measurement methods. A suitable thickness measurement device is available from Nanometrics (Milpitas, CA) or Nova Measuring Instruments (Phoenix, AZ). A computer may be utilized to calculate the optimal pad conditioning recipe based upon the measured film thickness and calculated removal rate, employing the models and algorithm provided according to the invention. A suitable integrated controller and polishing apparatus (Mirra with iAPC or Mirra Mesa with iAPC) is available from Applied Materials, California.

Exemplary semiconductor wafers that can be polished using the concepts discussed herein including, but are not limited to those made of silicon, tungsten, aluminum, copper, BPSG, USG, thermal oxide, silicon-related films, and low k dielectrics and mixtures thereof.

The invention may be practiced using any number of different types of conventional CMP polishing pads. There are numerous CMP polishing pads in the art which are generally made of urethane or other polymers. However, any pad that can be reconditioned can be evaluated and optimized using the method of the invention. Exemplary polishing pads include Epic™ polishing pads (Cabot Microelectronics Corporation, Aurora IL) and Rodel® IC1000, IC1010, IC1400 polishing pads (Rodel Corporation, Newark, DE), OXP series polishing pads (Sycamore Pad), Thomas West Pad 711, 813, 815, 815-Ultra, 817, 826, 828, 828-E1 (Thomas West).

Furthermore, any number of different types of slurry can be used in the methods of the invention. There are numerous CMP polishing slurries in the art, which are generally made to polish specific types of metals in semiconductor wafers. Exemplary slurries include Semi-Sperse® (available as Semi-Sperse® 12, Semi-Sperse® 25, Semi-Sperse® D7000, Semi-Sperse® D7100, Semi-Sperse® D7300, Semi-Sperse® P1000, Semi-Sperse® W2000, and Semi-Sperse® W2585) (Cabot Microelectronics Corporation, Aurora IL), Rodel ILD1300, Klebesol series, Elexsol , MSW1500, MSW2000 series, CUS series and PTS (Rodel).

In at least some embodiments, the method of the invention can be used to predict pad life for tool scheduling. For example, if the number of wafers, after each condition cycle decreases, one could predict a pad life end point and use that prediction to schedule retooling.

The present invention is described above under conditions where wafer polishing parameters are held constant. However, in at least some embodiments of the invention, the methodology can also be used together with an optimization engine when the wafer polishing parameters are changing through an optimization engine.

In at least some embodiments, pad conditioning optimization may be carried out together with optimization of wafer polishing. This can be accomplished through

optimization by having the optimization search engine's objective function minimize a function that describes both polishing and conditioning parameters.

Assuming n number of polishing parameters to be changed during the wafer polishing, N1, N2, N3....Nn, and y number of control parameters, Y1,Y2,. Yy, then

$$S = W_{N1}(N1_{previous} - N1_{current})^2 + W_{N2}(N2_{previous} - N2_{current})^2 + \ldots W_{Nn}(Nn_{previous} - Nn_{current})^2 + W_{\omega}(\omega_{previous} - \omega_{current})^2 + W_d(d_{previous} - d_{current})^2 + W_{Y1}(Y1_{previous} - Y1_{current})_2 + W_{Y2}(Y2_{previous} - Y2_{current})^2 + W_{Yy}(Yy_{previous} - Yy_{current})^2,$$

where W_x is a weighing factor for parameter x (e.g., N1, N2, Y1, Y1, F, etc.), ω is the pad rotational velocity, and d is the conditioning pad direction of spin. Other pad conditioning parameters can be included in the function. The optimization process then seeks to minimize S. Thus, the method of the present invention can be used under conditions when the polishing parameters are held constant or when the polishing parameters are to be changed through optimization.

Various aspects of the present invention that can be controlled by a computer, including computer or other controller **902**, can be (and/or be controlled by) any number of control/computer entities, including the one shown in FIG. 11. Referring to FIG. 11 a bus **1156** serves as the main information highway interconnecting the other components of system **1111**. CPU **1158** is the central processing unit of the system, performing calculations and logic operations required to execute the processes of embodiments of the present invention as well as other programs. Read only memory (ROM) **1160** and random access memory (RAM) **1162** constitute the main memory of the system. Disk controller **1164** interfaces one or more disk drives to the system bus **1156**. These disk drives are, for example, floppy disk drives **1170**, or CD ROM or DVD (digital video disks) drives **1166**, or internal or external hard drives **1168**. These various disk drives and disk controllers are optional devices.

A display interface **1172** interfaces display **1148** and permits information from the bus **1156** to be displayed on display **1148**. Display **1148** can be used in displaying a graphical user interface. Communications with external devices such as the other components of the system described above can occur utilizing, for example, communication

port **1174**. Optical fibers and/or electrical cables and/or conductors and/or optical communication (e.g., infrared, and the like) and/or wireless communication (e.g., radio frequency (RF), and the like) can be used as the transport medium between the external devices and communication port **1174**. Peripheral interface **1154** interfaces the keyboard **1150** and mouse **1152**, permitting input data to be transmitted to bus **1156**. In addition to these components, system **1111** also optionally includes an infrared transmitter and/or infrared receiver. Infrared transmitters are optionally utilized when the computer system is used in conjunction with one or more of the processing components/stations that transmits/receives data via infrared signal transmission. Instead of utilizing an infrared transmitter or infrared receiver, the computer system may also optionally use a low power radio transmitter **1180** and/or a low power radio receiver **1182**. The low power radio transmitter transmits the signal for reception by components of the production process, and receives signals from the components via the low power radio receiver. The low power radio transmitter and/or receiver are standard devices in industry.

Although system **1111** in FIG. 11 is illustrated having a single processor, a single hard disk drive and a single local memory, system **1111** is optionally suitably equipped with any multitude or combination of processors or storage devices. For example, system **1111** may be replaced by, or combined with, any suitable processing system operative in accordance with the principles of embodiments of the present invention, including sophisticated calculators, and hand-held, laptop/notebook, mini, mainframe and super computers, as well as processing system network combinations of the same.

FIG. 12 is an illustration of an exemplary computer readable memory medium **1284** utilizable for storing computer readable code or instructions. As one example, medium **1284** may be used with disk drives illustrated in FIG. 11. Typically, memory media such as floppy disks, or a CD ROM, or a digital video disk will contain, for example, a multi-byte locale for a single byte language and the program information for controlling the above system to enable the computer to perform the functions described herein. Alternatively, ROM **1160** and/or RAM **1162** illustrated in FIG. 11 can also be used to store the program information that is used to instruct the central processing unit **1158** to perform the operations associated with the instant processes. Other examples of suitable computer readable media for storing

information include magnetic, electronic, or optical (including holographic) storage, some combination thereof, etc. In addition, at least some embodiments of the present invention contemplate that the medium can be in the form of a transmission (e.g., digital or propagated signals).

In general, it should be emphasized that the various components of embodiments of the present invention can be implemented in hardware, software or a combination thereof. In such embodiments, the various components and steps would be implemented in hardware and/or software to perform the functions of the present invention. Any presently available or future developed computer software language and/or hardware components can be employed in such embodiments of the present invention. For example, at least some of the functionality mentioned above could be implemented using the C, C++, or any assembly language appropriate in view of the processor(s) being used. It could also be written in an interpretive environment such as Java and transported to multiple destinations to various users.

Although various embodiments that incorporate the teachings of the present invention have been shown and described in detail herein, those skilled in the art can readily devise many other varied embodiments that incorporate these teachings.